

Don't Expose Yourself: Discretionary Exposure to Political Information*

Gaurav Sood[†]

Yphtach Lelkes[‡]

November 19, 2017

*The paper benefited from comments by Marcel Garz, Stephen Goggin, and Benjamin Lyons. We are grateful to them. We are also grateful to Shanto Iyengar for letting us use data from surveys that we jointly fielded.

[†]Gaurav is a machine learning engineer. He can be reached at: gsood07@gmail.com

[‡]Yphtach is Assistant Professor of Political Science at University of Pennsylvania. He can be reached at: yphtach@gmail.com

Summary

The news media have been disrupted. Broadcasting has given way to narrowcasting, editorial control to control by ‘friends’ and personalization algorithms, and a few reputable producers to millions with shallower reputations. Today, not only is there a much broader variety of news, but there is also more of it. The news is also always on. And it is available almost everywhere. The search costs have come crashing down. So much so that much of the world’s information is at our fingertips. Google anything and the chances are that there will be multiple pages of relevant results.

Such a dramatic expansion of choice and access is generally considered a Pareto improvement. But the worry is that we have fashioned defeat from the bounty by choosing *badly*. The expansion in choice is blamed for both, increasing the ‘knowledge gap,’ the gap between how much politically interested and politically disinterested know about politics, and increasing partisan polarization. We reconsider the evidence for the claims. The claim about media’s role in rising knowledge gaps doesn’t need explaining because knowledge gaps aren’t increasing. For polarization, the story is nuanced. Whatever evidence exists suggests that the effect is modest. But measuring long-term effects of a rapidly changing media landscape is hard and may explain the results.

As we also find, even describing trends in basic explanatory variables is hard. Current measures are beset with five broad problems. First is conceptual errors. For instance, people frequently equate preference for information from partisan sources with a preference for congenial information. Second, survey measures of news consumption are heavily biased. Third, behavioral survey experimental measures are unreliable, and inapt for learning how much information of a particular kind people consume in their real lives. Fourth, measures based on passive observation of behavior only capture a small (likely biased) set of the total information consumed by people. Fourth, content is often coded crudely—broad judgments are made

about coarse units, eliding over important variation.

These measurement issues impede our ability to answer the extent to which people choose *badly*, and the attendant consequences of such. And improving measures will do much to advance our ability to answer important questions.

Keywords: Selective Exposure, Partisan Media, Echo Chambers, Filter Bubbles, Knowledge Gaps

We live in the proverbial information age. A myriad of stimuli constantly vie for our attention. And we constantly decide which stimulus to attend to. Much of the decision-making happens *automatically*, beyond our conscious control. Vital as such decisions are, the chapter doesn't discuss them. Instead, it focuses on the complementary question. How do people consciously choose what information to attend to?

We begin with a patent premise. People consume the media they prefer. They choose from a rich, near unlimited, menu. And deduce what they prefer based on information that is easily available to them. Their deductions are often flawed. Limited available information, finite cognitive capacity, a bias toward cognitive miserliness, among other things, circumscribe their ability to choose *correctly*.

On rare occasions what we want is not on the menu. More often, what we want isn't easy to find. Other times the available information leads us astray. For instance, when deciding, the name of the outlet is often one of the few pieces of information we have. But source cues can mislead. *The New York Times* (NYT) may lean liberal, but it often carries news that is congenial to the right (Garz et al. 2017; Barberá and Sood 2014). Other times, laziness explains what we consume. Merely changing the position of the channel on the menu affects what we consume (Martin and Yurukoglu 2014). These limitations mean that there is often a sizable gap between people's ideal points and the ideal point of what they consume.

However inexpertly and desultorily, people pick what they think maximizes their utility. Except what gives people utility in the moment often differs from what gives them utility after they have had a chance to reflect (Akerlof and Shiller 2015). In the moment, people often want to be entertained. On reflection, they often want to be improved. In the moment, they may pick burgers. On reflection, kale. We posit that for most people, most of the time, less considered goals likely dictate consumption.

Confining ourselves to decisions about consumption of political information, satiating interest in politics, interest in partisan (or ideologically) congenial information, and interest

in accurate information gives people a lot of utility. Aside from these, people have preferences about style. For instance, satirical over seriously delivered news or video over text. People also prefer learning about some issues more than others (see for e.g., [Iyengar et al. 2008](#)). For instance, old people likely prefer reading news about Medicare than student loans; immigrants, news about their home country to news from another country; the young, news about student loans than about Medicare. People also prefer negative information ([Trussler and Soroka 2014](#); [Arango-Kure, Garz and Rott 2014](#)).

But what people want depends on the context. For instance, even those who abstain from the news likely have eyes glued to it when there is a terrorist attack. By the same token, as the election day approaches, people consume more political information ([Gentzkow and Shapiro 2011](#); [Garz et al. 2017](#)). Preferences also evolve with time. For instance, our current affairs knowledge peaks around middle age (see [Appendix A](#)), suggesting that we pay the greatest attention to news then (see also [Garz \(2017\)](#)). Similarly, as we note above, which issues we find interesting depends on where we are in life.

Not only are preferences mutable, but they also vary across people. The taste for political information varies *immensely*. At one end are those who live and breathe politics, keeping track of every latest poll. At the other, much deeper, end are those who couldn't be bothered with knowing which party holds a majority in the House. (Only 47% of the people know the fact ([Bawn et al. 2012](#)).) Taste for congenial political information also varies tangibly (see, for e.g., [Iyengar and Hahn 2009](#)).

In toto, preference for political information is complex, variable, and mutable. Given the richness, we need to make some choices about what to explore in greater detail. Our interest in how people use the media is largely instrumental. We are interested in it to the extent that it affects more causally proximate variables in politics such as what people know about politics, how they feel about partisans of the main opposing party, incentives of politicians to communicate accurate information, etc. The rise of polarization and the potential rise in inequality of

political knowledge are concerning. And media's role in both is hotly debated. Thus, we opt for exploring preferences for political information and partisan congenial information.

We begin by considering theory and some evidence about preferences for political information and politically congenial information. We calibrate our understanding by looking at expected changes in key political variables of interest—political knowledge (and misinformation) and polarization, respectively. We then delve more deeply into the conceptual and methodological issues around measurement of consumption and preference for political information, politically congenial information, and the impact of changes in the media. Our in-depth exploration of measurement issues follows a set format. We define what we ideally want to measure. And then use the ideal to bring greater clarity to the consequences of compromises in measurement that various measures make.

A Preference for Entertainment

Political ignorance is in vogue again. A sharp rise in economic inequality in advanced democracies has brought to fore concerns about political ignorance. Many people think that inequality is a bad thing but still prefer policies that would increase it (Bartels 2005). Analysis by Bartels finds that ignorance explains a good bit of the puzzle. As goes for inequality, so goes for other political issues. At least some people would prefer other policies if they knew more (Luskin, Fishkin and Jowell 2002). For a fuller discussion on the importance of current affairs knowledge, see Graber (2004); Leighley (2003); Schudson (1998); Zaller (2003). Given its importance, it is distressing that people know so little about politics. For instance, in 2000, just 42% of the people knew that Bush wanted more restrictions on abortion than Gore (Bawn et al. 2012). Seven years after Putin first became president of Russia, 64% of the people couldn't identify him as the president of Russia (Kohut, Morin and Keeter 2007). What explains this umbra of political ignorance?

Like many things in life, you need a trifecta to learn about politics. You need opportunity. When given the opportunity, you need to be interested in capitalizing on the opportunity. And to capitalize on the opportunity, you must have the wherewithals. Or, access to political information, interest in consuming it, and the ability to process it explain what we know about politics (p. 335, [Luskin 1990](#)). Of the three, ability has likely increased over the past 50 years, though likely not enough to make much of a difference. This leaves us with interest and opportunity.

Interest in politics is often founded in differences in taste ([Luskin 1990](#); [Prior 2003](#); [2007](#); [Iyengar and Hahn 2008](#); [Prior 2003](#)). Taste in politics forms early and is stable through life ([Prior 2010](#)). Aside from taste, interest in politics is also shaped by its intrinsic and instrumental value. Some like to stay informed because it helps them vote ‘correctly’ ([Shineman 2012](#)). Others, because people they talk to value such information ([Sood 2014](#); [Genova and Greenberg 1979](#)).

Neither intrinsic nor instrumental value of political information is large enough for most people to consume much of it ([Hindman 2008](#); [Flaxman, Goel and Rao 2016](#)). For instance, over a three month period, only 4% of the users read 10 or more news articles and 2 or more opinion pieces on one of the browsers on their computer ([Flaxman, Goel and Rao 2016](#)). To give you more perspective, more than half the users viewed 991 or more pages during the three months on the browser.

Offline things are no different. Less than 8% of Americans tune into network news regularly, and less than 1% watch cable news regularly ([Prior 2009](#)). Rather than watch public affairs programming, Americans consume entertainment. In 2011, according to Nielsen, Americans spent 9.6% of the time spent watching television on watching the news. They spent 15.5% of the time watching television watching reality television, 11.4% watching sitcoms, 22.5% watching sports, and 41.1% watching drama.

Even those who prefer news over entertainment do not always choose news high in

nutritional value. Instead of news about major political issues, people often choose personality-centric, sensational news. According to Nielsen, in 1999, soft news was about as popular as network news (Prior 2003).

Unlike interest, the opportunity to consume news has grown considerably in the last few decades. Where there once was a piddly stream, today there is a biblical flood. Today, people have access to news 24 hours a day, on smartphones, tablets, and television. And there is a rich selection of style and viewpoints to choose from. But there is little evidence that people know more today (Kohut, Morin and Keeter 2007). (The only column in possible ascendance is misinformation, though even there it is hard to say much reliably.) Markus Prior posits an explanation for the stagnation: Today, not only is the opportunity to consume politics greater, so is the opportunity to escape it (Prior 2007).

Until a quarter a century or so ago, consumption of political news likely depended a shade less on how interested people were in politics. During certain times in a day, people could not easily escape political news. At night, when most people were at home and seated in front of their televisions, there was no other programming on the television except for news. And rather than turn off the television and wait for the next sitcom, some people sat through the news (Robinson 1976). The *captive* audience was likely not very interested in news, but people likely absorbed some information (Keeter and Wilson 1986; Krugman and Hartley 1970; Zukin and Snyder 1984).

But the evidence for broadcast television's impact on political knowledge points the other way. Exploiting exogenous variation in the introduction of television, Gentzkow (2006) find that television led to a 25–50% decline in turnout. They attribute the decline to lower informativeness of television vis-à-vis newspapers. Thus, despite creating an inadvertent audience for news, broadcast media reduced net consumption of public affairs.¹

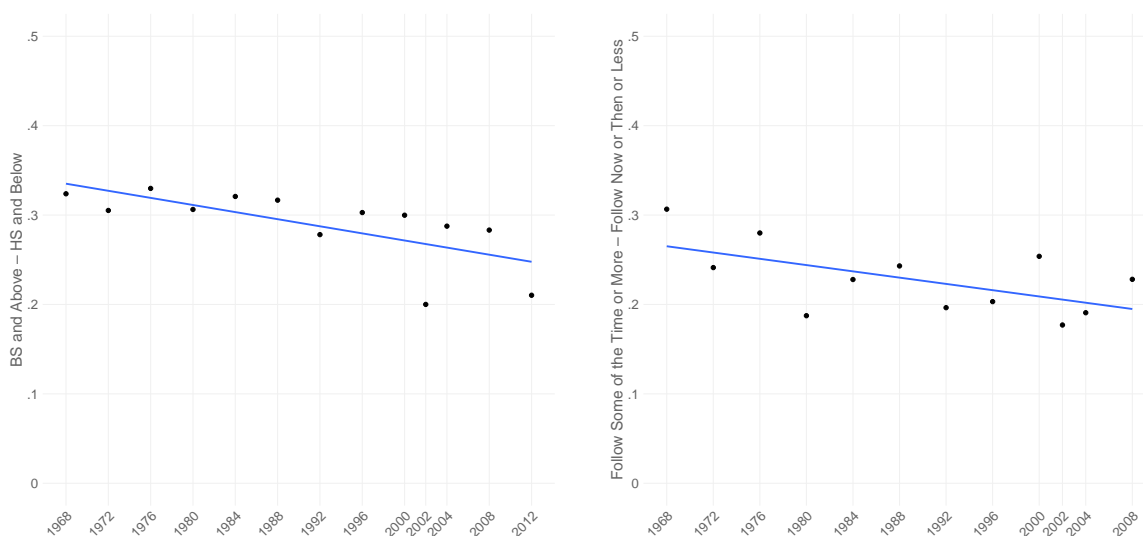
¹This effect is contingent on media systems, however. Sørensen (2016) finds that television increased turnout in Norway, where television is more public oriented than in the United States.

But what about cable, Internet, and mobile? Are the politically disinterested taking advantage of the greater opportunity to opt out of the news? If they are, the knowledge gap between those less and more interested in politics should be increasing. To investigate, we turn to data from the American National Election Studies (ANES). The ANES asks interviewers to rate “respondent’s general level of information about politics and public.” The interviewers rate respondents on a 5-point scale ranging from “Very Low” to “Very High.” These ratings are a good measure of political knowledge (pg. 333-344, Zaller 1992). ANES also includes measures of political interest and education. (See Appendix A for question text and variable names.) We recoded all the responses linearly to lie between 0 and 1.

To check how the knowledge gap between politically interested and disinterested, and the better-educated and worse-educated has evolved, we began by estimating year-wise bivariate regressions between political knowledge and political interest and education. The trend in the coefficients—correlation with year—gives us a bare-bones estimator of the trend in the knowledge gap. Contrary to expectations, for both political interest and education (coded in 7 categories), the correlations are robustly negative ($r_{\text{polint}} = -.43$, $r_{\text{education}} = -.51$). Regressing knowledge on interviewee characteristics (age, age squared, gender, race) and mode of the interview and interaction between political interest and year of the interview (recoded to lie between 0 and 1) yields an estimate of $-.06$ for the interaction term, suggesting a declining knowledge gap. Replicating the same regression with education yields a coefficient of $-.11$ for the interaction. Controlling for interviewer race and gender leads to a loss of a majority of the sample, but the magnitude of the coefficient on the interaction term becomes larger. The upshot is: contrary to expectations, the knowledge gap between the better-educated and the less well-educated is declining modestly. To illustrate the point, we plot the knowledge gap between those with college or more and those with high school or less (note the y-axis) (see Figure 1).

Both political interest and political knowledge are measured much too crudely to rule

Figure 1: Knowledge Gaps Across Education and Political Interest



out anything confidently. But the data suggest skepticism about the new conventional wisdom about the new salience of political interest. What is likelier is that changing information landscape has made some people more knowledgeable. This is in line with recent work showing that as people spend more time online, they consume more of both news and entertainment (Lelkes, Sood and Iyengar 2015; Lelkes 2017). The rapid decline in local news outlets has also likely affected how much people know about local politics (see also Snyder Jr and Strömberg 2008). But many fundamental questions about how the transition to new media has affected what and how much people know about politics remain. And one key reason the questions remain unanswered is because of the enormous challenge of measuring consumption well. We discuss these challenges later in the chapter.

A Preference for Congenial Political Information

Some people prefer congenial information because they think it is more accurate (Gentzkow and Shapiro 2006; Lord, Ross and Lepper 1979; Metzger, Hartsell and Flanagin 2015). It is tempting to label judgments of greater accuracy as wishful thinking. But there are good

reasons to resist the temptation. Often people do not have much information on which to judge the accuracy of a piece of information. Often enough, they also lack the skills to deduce how accurate a piece of information is. Deductions from study design and sample size remain beyond the pale even for scientists. For instance, a vast majority of scientists obey the law of *small* numbers (Tversky and Kahneman 1971). Bereft of the two tools, people often have no option other than to judge the credibility of a number based on how far it is from their priors. The further it is, the less credible. This is the rule we use when struck with vertigo, we discount the possibility that the roof is spinning. The fact that congenial information is seen as more accurate has consequences beyond consumption. It also means that congenial information likely influences opinions more—it makes sense to take greater account of more trustworthy information.

Others prefer congenial information because it is unpleasant to receive uncongenial news (Festinger 1957). Uncongenial news is unpleasant because it shows us to be wrong. And some people prefer to be proven right than be correct. There is a potential further wrinkle to psychic rewards for congeniality. People may feel ‘losses’ from uncongenial information more than rewards from congenial information (Garrett 2009), though see Garz et al. (2017); Sears and Freedman (1967; 1963).

Finally, others may encounter congenial information as a function of their social context. Those living in ideologically homogeneous information environments are more likely to be exposed to the prevalent ideology of their environment Sears and Freedman (1967; 1963). If, for instance, most of your online social network is liberal, the news you come across in your social network feed is more likely to be liberal (Bakshy, Messing and Adamic 2015).

The evidence for these theories is often hard to muster. And the difficulty is often compounded by a common confusion. People often confuse preference for information from a congenial source with a preference for congenial information. The two are not the same. First, prominent sources with partisan reputations often carry lots of uncongenial information (Garz

et al. 2017; Barberá and Sood 2014). Second, preference for news from a congenial source may not stem from perceived accuracy and psychic benefits alone. News from a congenial source can also provide greater instrumental value. Advice by an agent who shares your preferences is more valuable (Chan and Suen 2008).

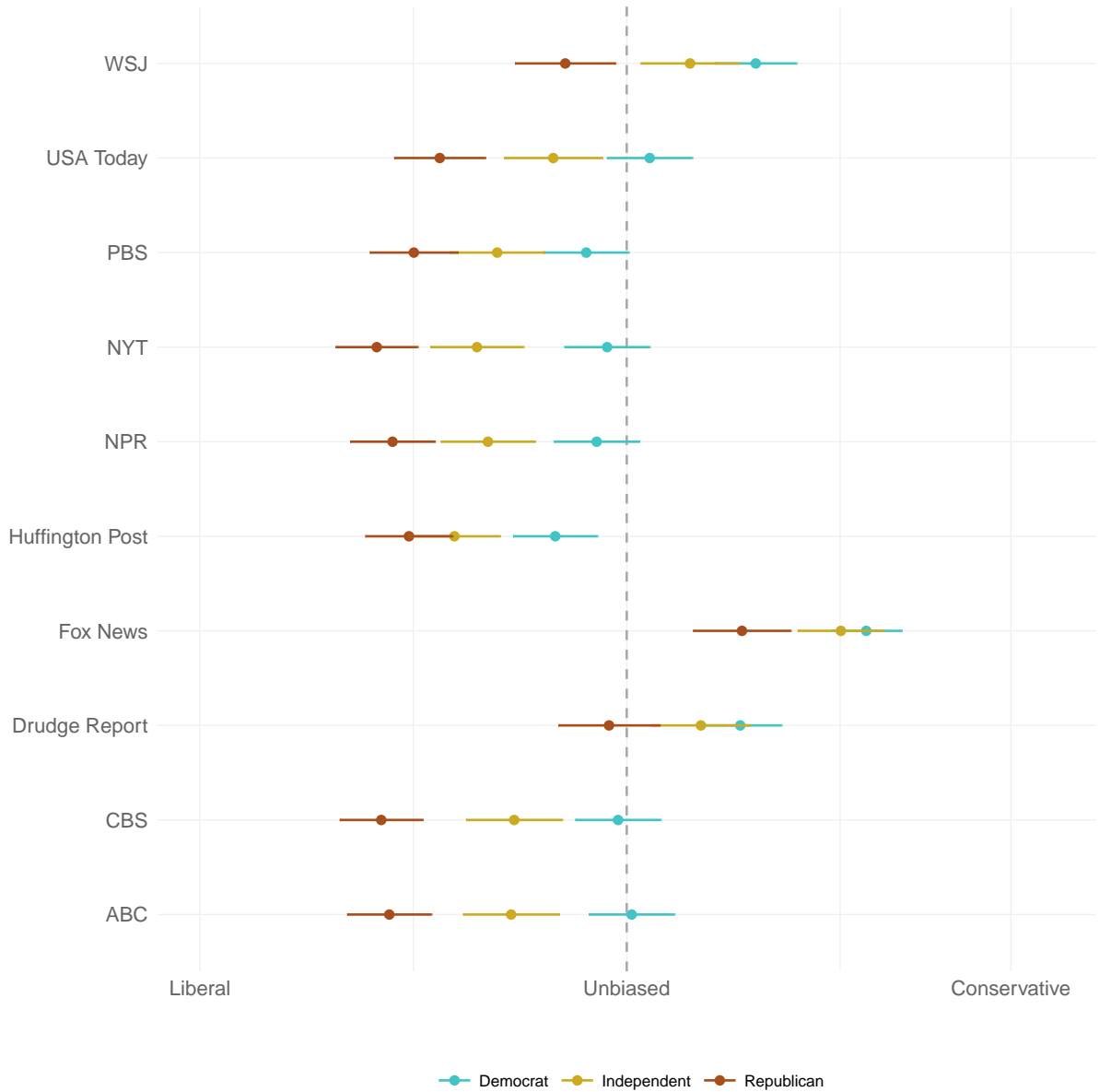
If we ignore the instrumental value of congenial sources, we are left with perceived accuracy and psychic benefits. But the point about perceived accuracy is still underappreciated. So to complement other work showing people trust congenial sources more (see, e.g., Metzger, Hartsell and Flanagin 2015), we conducted a survey. Along with Shanto Iyengar, in 2011, we surveyed 1,000 respondents recruited by YouGov. We asked the respondents to place various news organizations on a 7-point ideological scale. The scale ranged from liberal (1) to conservative (7), with the midpoint labeled as ‘no bias at all.’ Objective assessments put NYT, NPR, ABC, CBS, USA Today, Huffington Post, and PBS to the left of center. And Fox News, *The Wall Street Journal* (WSJ), and the Drudge Report to the right of center (Barberá and Sood 2014).

Partisans rated outlets that shared their outlook as less biased than independents or opposing partisans (see Figure 2). Democrats on average rated NYT, NPR, and PBS near the midpoint. Republicans thought Fox News, WSJ, and the Drudge Report were less biased than other outlets. And Independents’ ratings fell between Democrats’ and Republicans’. All this was expected. But there was one surprising result. Independents, on average, got the sidedness of the slant correct for all outlets (see, Barberá and Sood 2014).

To parse the extent to which people choose news from a congenial source because they think it is more accurate from because they think it is more pleasant, we need to measure both. We need to measure the ideological distance between a person’s ideal point and the ideal point of the outlet. And we need a measure of perceived accuracy. We can then regress the choice on ideological distance and perceived accuracy. That perceived accuracy and perceived ideological distance are likely endogenous is immaterial. Perceptions are the causal quantities of interest. But even this scheme assumes that we know the utility function. We don’t.

There is another, simpler, way of estimating the extent to which the news selection is ideological than driven by trust. We can exploit the fact that judgments about trust and accuracy are generally made about media sources than about news stories. Thus, we can build a measure using revealed preferences over a series of choice tasks in which the congeniality

Figure 2: Perceived Bias of News Organizations by Party Identification



of the stories and the sources to which they are attributed to is manipulated. The key thing is that we need to limit inference to within outlets.

Regardless of whether people prefer partisan sources because they trust them more or because they expect psychic rewards, people tend not to consume a lot of information from such sources (Prior 2009; Gentzkow and Shapiro 2011; Guess 2016). But underneath the modest average hides sizable variation. Strength of partisanship is correlated with preference for information from congenial sources (Garrett 2009; Iyengar and Hahn 2008) (see also Appendix B.2). (We expect stronger partisans to also consume more news from partisan sources.)

There is a further nuance to the point. The correlation between preference for information from congenial sources and strength of partisanship may be becoming more modest. Some survey evidence points to the tautology that the politically disinterested sometimes opt for entertainment over news (Prior 2005; 2007; Arceneaux, Johnson and Murphy 2012). But increasingly, choosing entertainment doesn't mean forgoing partisan news outlets. Partisan news outlets, undoubtedly aware of the demand for apolitical news, now carry a lot of it. And this gives particular teeth to the evidence that people also prefer consuming *soft* news from partisan congenial sources (Iyengar and Hahn 2008). Consuming apolitical news from a partisan source likely increases the chances of being exposed to political headlines and imagery that is politically congenial.

To shed more light on the issue, we fielded a new survey. In 2013, we surveyed 2,000 people recruited by YouGov. We showed each respondent four screens with six news stories each. On each screen, we asked them to pick the story they were most inclined to read. For cover, we told them that we were developing a new way of displaying online news. In keeping with the cover, the screen layout followed that of Google News. A Google News logo also appeared in the upper left corner. The story headlines and logo of the sources were split into two rows of three headlines. (See Appendix B.1 for screenshots of the selection screens.) To account for the impact of position, we randomized the position of stories.

Like [Iyengar and Hahn \(2008\)](#), we randomly assigned identical hard and soft news stories to different sources. Of the 24 stories, 12 were political and 12 apolitical. Political stories were on electoral politics and the economy. Apolitical stories focused on celebrity life, sports, movies, and sleep disorders. The first and third trial showed logos of news organizations, e.g., Fox News. The second and the fourth trials matched stories included information about the channel and the show, e.g., The O'Reilly Factor (Fox). All four screens featured two left-leaning sources, two right-leaning sources, and two non-partisan sources.

Partisans choose apolitical stories from partisan congenial sources about as often as they choose political stories (see [Figure 3](#)). Both Republicans and Democrats picked political and apolitical news from congenial outlets 10% more often than from uncongenial outlets.

Republicans, however, prefer information from congenial sources more than Democrats (see also, [Garrett 2009](#); [Iyengar and Hahn 2008](#)). Republicans chose right-leaning sources twice as often as left-leaning sources—50% versus 22%. But this is not conclusive proof that Republicans prefer information from congenial sources more than Democrats more generally. First, all the studies treat MSNBC and Fox interchangeably. Fox is taken to be as right of center as MSNBC is to the left. They aren't ([Barberá and Sood 2014](#); [Martin and Yurukoglu 2014](#)). Second, Fox has been right of center for a much longer time than MSNBC has been left of center ([Martin and Yurukoglu 2014](#)). Thus, Fox News is likely to have greater 'brand' recognition than MSNBC.

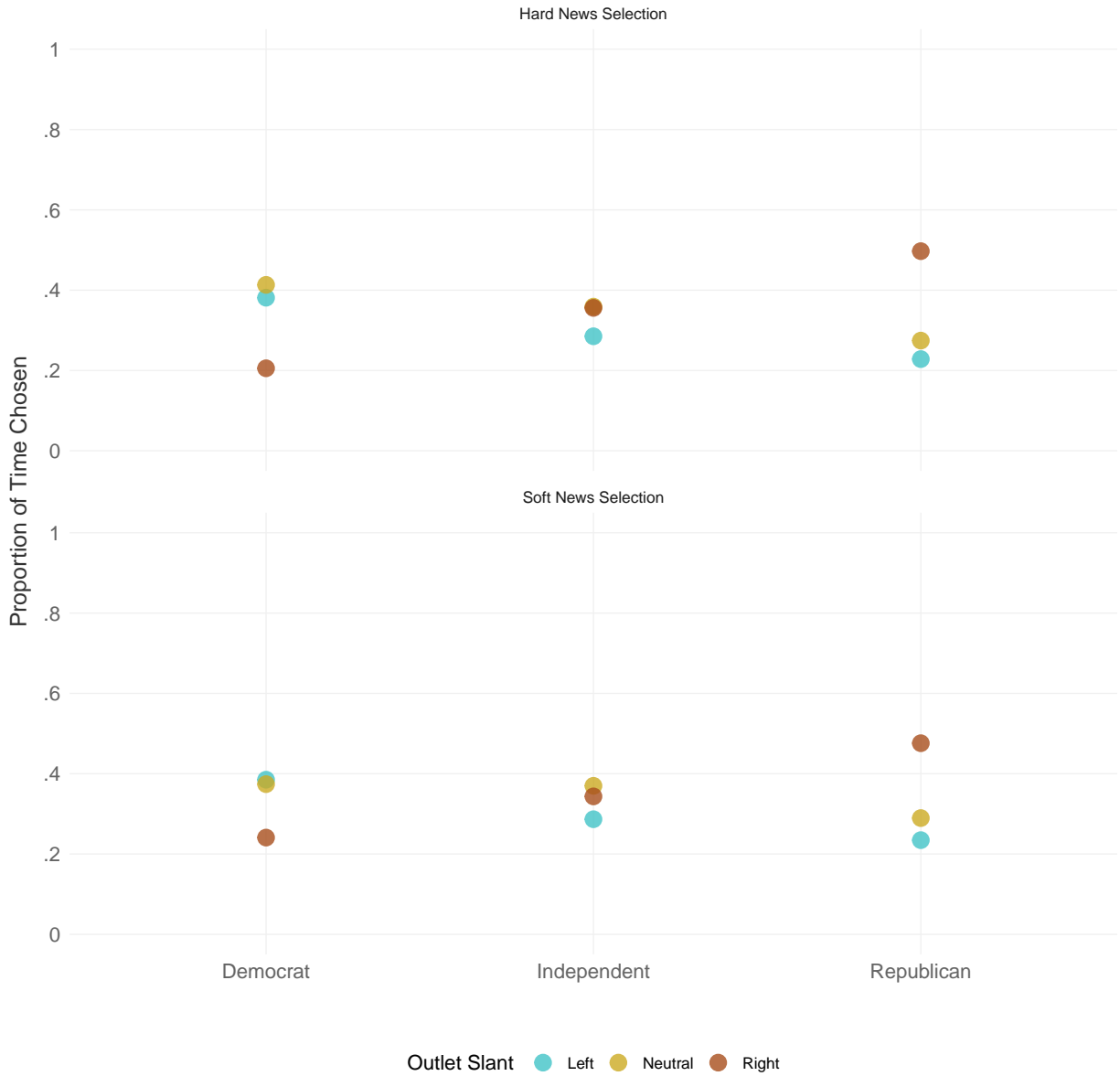
Even if people consume all the information from partisan sources, it does not mean that it will have a large effect on their attitudes and behaviors. Given the polarized sort into partisan media, the distance between positions of partisan media and the consumer may be small ([Bennett and Iyengar 2008](#)). And consumers may be aware of the biases, and may appropriately down weight what they hear from different media sources.

Neither reason is compelling. Many media organizations are more ideologically extreme than the median partisan ([Barberá and Sood 2014](#)). Second, people are generally bad at

estimating bias (see Figure 2).

Thus, we expect consumption of congenial over uncongenial bits of information to affect politically important variables. It should affect partisan affect, what people know, policy positions, trust, political participation (donating, attend a rally, voting, etc.), and whom they

Figure 3: Patterns of News Selection by Slant of Source



vote for (e.g., [Stroud 2010](#); [Levendusky 2013](#); [Arceneaux, Johnson and Murphy 2012](#); [Melican and Dixon 2008](#); [Martin and Yurukoglu 2014](#)).² Existing stores of knowledge likely condition the extent to which consumption of partisan media affects various variables ([Zaller 1992](#)). The ignorant may be swayed by small changes in consumption. The bare stores of political knowledge of an average American likely explain why small changes in their media diet matter ([DellaVigna and Kaplan 2007](#); [Martin and Yurukoglu 2014](#)).

Exposure to uncongenial information may also polarize. A Democrat watching Fox News Channel (FNC) may think that its coverage is defamatory and hateful. And that people who typically watch FNC endorse this. And may revise their views of Republicans as a result.

Hitherto, we have assumed that the effects of media are limited to the people who consume it. But that is almost certainly wrong. There is a long literature in Communication on the two-step flow ([Katz 1957](#)). And in an era where technology for sharing information has never been easier, it is likely that the consumption also affects the attitudes and preferences of people who interact with people who consume partisan media. In one recent study, [Druckman, Levendusky and McLain \(2015\)](#) exposed random subsets of participants to partisan news and then grouped all subjects in randomly composed discussion groups. They find that the effects of partisan news had an effect on the attitudes of those who were not exposed to partisan media.

In all, the evidence until now suggests that people do not consume a whole of partisan information, though strong partisans likely consume more of it. The effect of the amount of partisan information that a person consumes and their political attitudes is also modest ([Lelkes, Sood and Iyengar 2015](#)), though it may be because of issues in measurement, which we turn to next.

²In [Appendix B.2](#), we provide some further evidence from the YouGov study which we discuss earlier in the section about the correlation between preference for news from partisan sources and what people know. People who prefer news from conservative outlets knew more about the IRS scandal than who preferred news from liberal or neutral outlets.

Measurement

Both preferences and behavior are worth studying. But often we have access to nothing more than imprecise measures of behavior, be they self-reports or passive observations of consumption through some narrow porthole, e.g., data from the Bing toolbar (Flaxman, Goel and Rao 2016). Inferring preferences from such data is tricky. And it is trickier still when the number of devices and software platforms on which people consume keeps increasing.

Letting t denote all information consumed, and p denote all political information consumed, the conventional estimand for preference for and consumption of public affairs is $\frac{p}{t}$. Letting c denote total congenial political information consumed, for ‘partisan selectivity’, it is $\frac{c}{p}$. As a result of Garrett 2009, another corresponding ratio was added. Letting u denote total uncongenial information consumed $\frac{u}{p}$.

Aside from the usual caveats that go into inferring preferences from consumption, it isn’t clear if the ratios are appropriate conceptualizations for all dependent variables. If the dependent variable is partisan affect, how ‘selective’ one is may not matter as much as the net imbalance in consumption—the difference between the number of congenial and uncongenial bits consumed (Lelkes, Sood and Iyengar 2015). For instance, someone who consumes five conservative units and ten liberal units of information is by the ratio measure as selective as someone who consumes 50 conservative units and 100 liberal units. If the two hypothetical people start with the same partisan affect and knowledge, the impact of consuming 50 more liberal units of information is very likely different than the effect of consuming five more liberal units of information.

A yet more unequivocal case can be made for political knowledge. If we want to predict political knowledge (after regressing out political interest), surely, it is the total bits of political information consumed that is a more appropriate right-hand side variable than the proportion of media consumption that is news. More generally, the trouble with normalizing is that we

assume people consume similar amounts of total (political) information. They don't. To our point, the Finns and the Danes know *more* about *both* celebrity affairs and public affairs than their US counterparts (Curran et al. 2009). And people living in areas with faster Internet consume more news (Lelkes, Sood and Iyengar 2015; Lelkes 2017).

Whatever the estimand, to estimate it, we need to make a series of compromises. And we need to keep track of the various ways inferences are affected by the decisions we make. Ideally, we would like to track all the information consumed, precisely measure ideology of each unit of information on the same scale as the ideology of consumer. We could then look at the amount of, and the distribution of the ideology of the information consumed. If we so wanted, we could also summarize various moments of the distribution of ideological distances for each person. And if we were worried about the dimensionality of ideology, we could do the same by issue area. But such granular data are hard, if not impossible, to get. And given the limitations, scholars have made a host of simplifying assumptions about each of the components of the measure. For a more systematic look at assumptions behind current measures, we deal with assumptions about the measurement of ideology, and consumption independently. Comparing the measures we have to measures we would like to have illuminates how the current measures fall short.

Measuring Consumption

In this chapter, our focus is on consumption of varieties of political information. The genus is political information. And the species of this genus differ in congeniality, among other things. But what is political information? All information that influences people's political attitudes or behaviors? If so, then limiting ourselves to news is likely too constraining. Popular television shows like *The Handmaid's Tale*, *Narcos*, and *Law and Order* have clear political themes. And watching them likely affects how people think about politics. For instance, watching *Law and Order* may convince people that public prosecutors are just and efficient (Sood and Trielli

2016). Shows like *Will and Grace* and *The Cosby Show* may be less clearly political, but they also have a political subtext. And more time people spend living in the social worlds created by these shows, the more likely they are to adopt its beliefs (Shanahan and Morgan 1999).

To shed more light on the political content of entertainment shows, we used data from the 2008 NAES. The survey asked respondents whether they watched various shows. The survey also asked respondents about their partisanship. Like Dilliplane (2011), we used partisanship of self-reported audiences to infer the ideological content of the show. Letting t_D denote the number of Democrats reporting watching the show, t_R the number of Republicans watching the show, we estimated $\frac{t_D}{(t_D+t_R)}$ for each show. For comparison, we present scores for The O'Reilly Factor and The Daily Show along with prominent entertainment shows. The partisan composition of the audience varies a lot (see Figure 4). And much of the variation is expected—audiences for *24* and *CSI Miami* are less Democratic than audiences for *The Ellen DeGeneres Show* and *Oprah*. This variation very plausibly partly reflects the ideological content of the shows. But the key point is not that this is a good measure of the politics of the show. But the point is that we may want to take account of exposure to 'entertainment shows.'

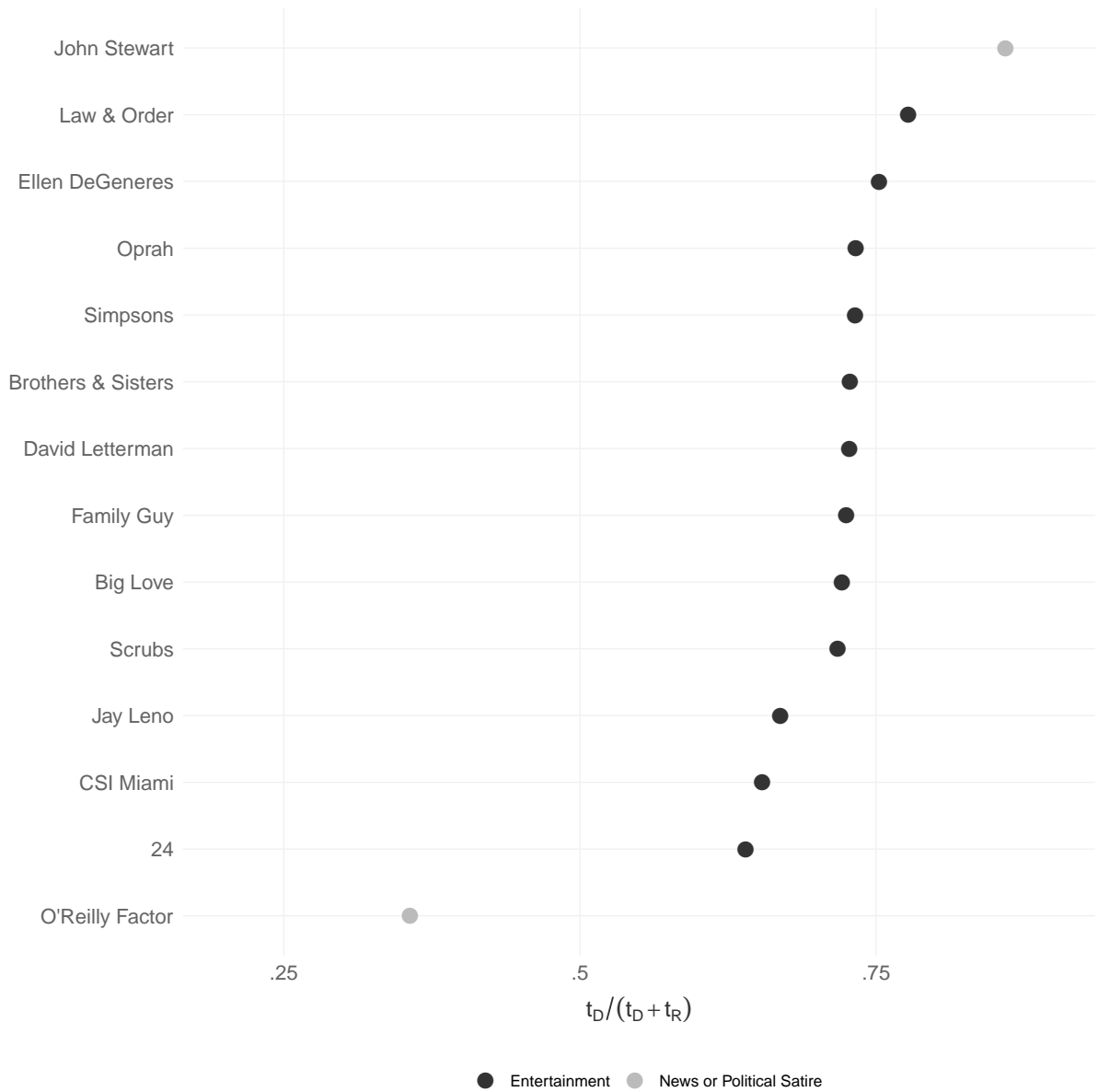
Even if we limit ourselves to news, the domain is still not clear. Is news about a bank robbery relevant political information? What about Hillary Clinton's haircut? To the extent that each of these affect people's attitudes, they are arguably pertinent. But care is needed to define the domain.

When questions about the domain are settled, we can get on with measurement. We can measure media consumption by asking people what media they consume, and by tracking what media they actually consume. Asking people is common. But tracking consumption is becoming more common. Both have their challenges.

Passive Observation of Consumption

Measuring consumption via passive observation may seem without its faults, but it isn't. First, passive collection often needs consent. This means people are aware that they are sharing their

Figure 4: Partisanship of Audiences of Entertainment Shows



Note: Data are from the 2008 NAES.

data with someone. And they may change their behavior in response to that (Rosenthal and Rosnow 2009). If that concern is wished away, another concern needs exorcising. Often, we only get to monitor consumption on a single application on a single device (see Flaxman, Goel and Rao 2016; Guess 2016). But people use many applications on many devices. Additionally, some panels will log time on a website if a tab is left open on the browser, while the media consumer is on a different tab. And using behavior from one application on one device to infer preferences and consumptions requires making bold assumptions or additional data. For instance, to estimate preferences, we must assume that the proportions are the same across. To estimate consumption, we need data on how often people use different devices. Finally, the sample for these panels is often quite biased. For instance, studies using tracking data from Internet Explorer and Bing (Flaxman, Goel and Rao 2016), which accounts for 22 percent of market share, might yield qualitatively different results from a study that tracks Google users, which accounts for over 60 percent of the search market.

Yet more serious issues attach to the measurement of what people are consuming. To infer what people are consuming online, we can rely on ad hoc lists of domains. For instance, Flaxman, Goel and Rao (2016),

...select an initial universe of news outlets (i.e., web domains) via the Open Directory Project (ODP, dmoz.org), a collective of tens of thousands of editors who hand-label websites into a classification hierarchy. This gives 7,923 distinct domains labeled as news, politics/news, politics/media, and regional/news. Since the vast majority of these news sites receive relatively little traffic, to simplify our analysis, we restrict to the one hundred domains that attracted the largest number of unique visitors from our sample of toolbar users. This list of popular news sites includes every major national news source, well-known blogs and many regional dailies, and collectively accounts for over 98% of page views of news sites in the full ODP list (as estimated via our toolbar sample)

Using such lists to measure political content raises three concerns. First, there is the danger of treating a non-news site as a news site. Second, there is the danger of missing some news sites. And third, there is the danger of assuming that news sites only carry news about politics.

Lists like those published by DMOZ seem well-enough curated not to contain too many false-positives. The key question is about how to calibrate the false negatives. Here's one proposal. Take a large random sample of the browsing data. Create a list of unique domains in the data. Then use a free large reputable database like Shallalist to get the kind of content hosted by the domain. For the domains that are not in the database, query a reputable web service like Trusted Source that provides the kind of content hosted by the domain. You can query all the domains using a web service, but the number of unique domains in the browsing data can easily run into millions. Comparing against Shallalist first helps reduce the amount of querying. Use the results to estimate the domain names carrying news that are not in the DMOZ list. Also, estimate missed visitation time and missed page views. The downside to relying on commercial vendors is that we do not know or control how they measure. Relying on replicable methods to code the content of the domains is best.

To address the issue of apolitical content on news sites, we can exploit the URL structure. Many news sites include semantic information in their URLs. For instance, a sports story will often have a URL with “/sports/” in it. [Flaxman, Goel and Rao \(2016\)](#); [Bakshy, Messing and Adamic \(2015\)](#) use this trick to categorize content. And they estimate false positive and negative rates by manually coding a small sample of article text.

Surveys

The problems with survey data can be broken down into issues with instrumentation and issues with respondents.

Survey instruments are often blunted by two kinds of sampling biases. The first is the

overrepresentation of salient options. Even the list in [Dilliplane, Goldman and Mutz \(2013\)](#) has but a few of the most popular shows on television. Such a skew will yield biased estimates if people are more (or less) selective about less mainstream political sources. For instance, if people went to mainstream news sources to get the information, and to less mainstream sites carrying sympathetic interpretation of facts for an opinion about the facts, instruments with salient sources would underestimate partisan selectivity. However, given the huge skew in consumption ([Gentzkow and Shapiro 2011](#); [Guess 2016](#); [Hindman 2008](#)), the consequences of such bias may not be severe.

A graver sampling problem on instruments pertains to the ideological composition of the domain. Often news media consumption batteries include equal proportions of conservative, liberal, and moderate news sources (see, for e.g. [Iyengar and Hahn 2008](#); [Knobloch-Westerwick and Meng 2011](#); [2009](#)) . The true composition of the domain is likely closer to a bulk of the choices being at or near the center, with only a few outlets with clear ideological slants. To credibly infer preferences from choices from such a biased sample is hard. And estimates of the proportion of time people watch news from partisan sources are near impossible to make with such data.

Moving to respondents, many respondents give biased responses on surveys. In the 2008 NAES, 75% of Americans reported that they read the news at least daily in some form. The number is hard to reconcile with another number we state above: 43%—the percentage of people on a national survey who did not know which party is more conservative. The answer to the puzzle is that survey measures of political news consumption are grossly inflated. A comparison between survey estimates of the audience for network news and passive observation measures suggests inflation of nearly 300% ([Prior 2009](#)).

Another kind of bias in responses is expressive responding ([Prior 2013](#)). People are allegedly particularly prone to overstating the extent to which they watch congenial media. Republicans may report watching Fox News because they want to signal that they consume

news that is consistent with their self-identity. By the same token, Democrats may be averse to acknowledging that they watch Fox News. This potentially explains why estimates of consumption of partisan media uncovered by survey data are much larger than obtained by passively collected consumption data.

Given the problems with self-reports, survey instruments that rely on behavioral measures are plausibly better (see [Iyengar and Hahn 2008](#)). In general, these behavioral measures are structured as follows. A respondent sees a few screens with a few headlines each. The source to which each of the headlines is attributed to is randomized. And a tally of the extent to which people click on stories from a congenial outlet over others serves as an estimate of the preference for information from congenial sources. (Though it is often misinterpreted as a preference for congenial information.) The behavioral measure suffers as much from sampling problems as conventional survey measures. For another, when you have a small choice set, the task of choosing is fundamentally different from the real world. Lastly, because of a small number of trials, the measure also tends to be noisy.

To shed more light on issues with behavioral measures, we use data from the same YouGov study we discuss in Section (also see [Appendix B.1](#)). We gauged the reliability of measures of preferences for partisan congenial news sources using a small number of trials. Like [Iyengar and Hahn \(2008\)](#), we randomly attributed the same political and apolitical news stories to different sources and asked participants to select the story they were most inclined to read. To estimate the reliability, we calculated the correlation between trials. We coded congeniality trichotomously: congenial, neutral, or uncongenial. The correlations between trials are alarmingly low. The polychoric correlation between any two trials range between .06 to .20. And the correlation between choosing political news in any two trials is between -.01 and .05.

To probe the validity, we calculated correlations with indicator variables. The correlation between preference for congenial news and preference for political news is extremely

weak ($r = -.05$). (See [Appendix B.1](#) for the operationalization of these measures.) This is puzzling. We expect those who prefer political news to prefer politically congenial news more [Iyengar et al. \(2008\)](#). In fact, self-reported news interest, while more strongly correlated with a preference for congenial news than the latent measure ($r = .21$), was still only moderately correlated with the latent trait model based on manifest choices ($r = -.36$). Similarly, the correlation between political knowledge and preference for congenial news was $r = .21$, while the correlation between political knowledge and preference for soft news was $r = -.36$. In fact, the correlation between political knowledge and self-reported interest in news was much stronger ($r = .59$). This suggests that latent trait models based on a small number of choices are less valid than self-reports.

Measuring Ideology

The modal study on selective exposure makes three decisions about the measurement of ideology. First, it judges ideology at the outlet level. Making judgments about the ideology of Fox News than say *Fox and Friends*. Second, it takes ideology of the outlets as self-evident, skipping formal measurement altogether. And third, it makes categorical judgments about ideology. It judges media as liberal (Democratic-leaning) or conservative (Republican-leaning). See for instance, [Iyengar and Hahn 2008](#); [Stroud 2010](#); [Levendusky 2013](#); [Garrett 2009](#); [Knobloch-Westerwick and Meng 2009](#). Each of the decisions has its attendant problems.

Making judgments at the level of outlet means measuring the first moment but not any higher moments. There is mean, but no variation. Or skew. Coding all the bits of information as having the same ideology seems unwise. The variance in ideological positions within outlets is sizable ([Barberá and Sood 2014](#)). A column by David Brooks is often poles apart from a column by Paul Krugman. To illustrate how measuring at the outlet level can mislead, let's consider a toy example. Say a Republican only visits <http://foxnews.com> and <http://nytimes.com>, visiting each 10 times. The conventional estimand for selective exposure

is the number of visits to Fox News divided by total site visits. In this case, we would arrive at an estimate of $\frac{10}{20}$. If Fox News had seven conservative and three liberal stories and NYT had five liberal and five conservative stories, the correct estimate is $\frac{7}{20}$.

The second point is self-explanatory. Measuring ideology by fiat is unscientific. Any claims based on such measures are suspect.

The third point deserves greater attention. By definition, coarse judgments do not capture 'enough' information. And the uncaptured information can be vital. To illustrate the point, consider the following. Say we only code a piece of information as Republican-leaning or Democratic-leaning. Assume also that a Republican consumes a bit of Republican-leaning information. We will infer that as a Republican consumed a bit of congenial information. Say on a finer 1 to 7 liberal to conservative scale, the consumer is at 6 and bit of information is at 5. Knowing that, we can say the Republican consumed something more liberal than his own ideological position. If the bit was at 7, we would say the consumer consumed something to the right of where they are. Crude measurement means treating consuming a bit of information at 5 and another at 7 the same.

Aside from concerns about subjectivity and crudeness of measures, there are questions about how to measure ideology. There are three broad ways we can get at the ideology of content. We can measure the ideology by asking people about their opinions, by tallying the ideology of the audience, and by coding the content. There are a lot of ways of doing each. And each way has its issues.

One way to measure slant is to ask consumers to rate the ideology. For instance, [Dilliplane \(2011\)](#) uses data from the 2008 National Annenberg Election Study (NAES). The survey asked respondents which outlets they used for information about the campaign. And then asked consumers of each outlet which presidential candidate it favored. If at least 25% of the self-reported consumers rated the outlet as favoring the Democratic candidate, [Dilliplane \(2011\)](#) rated it as left-leaning. Switch the candidates, and you have the version for coding an

outlet as right-leaning. All the remaining outlets were coded as neutral.

The audience rating based measure has three big weaknesses. First, the measure is subjective. People are bad monitors of bias, with estimates of bias correlated with congeniality (for e.g., [Morris 2007](#); [Vallone, Ross and Lepper 1985](#)). [Dilliplane \(2011\)](#) uses ratings from Independents to try to address this concern. But Independents are not ideological moderates. Second, the rationale for the 25% threshold is unclear. Third, the judgments are coarse. For instance, both *Good Morning America* (GMA) and *Countdown with Keith Olbermann* (CKO) are coded as left of center. This means coding someone only watching CKO as consuming as much congenial information as someone only watching GMA.

Another way to measure the slant of content is to use audience's ideology as proxy ([Van Kempen 2007](#); [Goldman and Mutz 2011](#); [Flaxman, Goel and Rao 2016](#)). For instance, [Flaxman, Goel and Rao \(2016\)](#) label an outlet conservative if more than half its audience supports the Republican candidate. But well-tread concerns about inferring ideology from behavior apply. For instance, New York City (NYC) is heavily Democratic. And NYT is the city's local newspaper. Say people in NYC read NYT because they like to get local news, not because it is liberal. In such cases, inferring that NYT is liberal based on the reading habits of NYC residents will be ill-advised. One way to mitigate these concerns is to specify a more complete utility function. And estimate relative ideological location as functions of key aspects of information that are thought to give utility to people ([Gentzkow and Shapiro 2011](#); [Barberá and Sood 2014](#)). The functional form is open to challenge, and so is what is included in the function, though calibration of structural parameters can be done with experiments ([DellaVigna, List and Malmendier 2012](#)).

Another way to measure ideology is to code content ([Gentzkow and Shapiro 2010](#); [Groseclose and Milyo 2005](#); [Ho, Quinn et al. 2008](#), see for instance,). But what aspect of content to code? One aspect of content that is informative about ideology is who is cited. [Groseclose and Milyo \(2005\)](#), for instance, track the number of times media outlets and law-

makers cite liberal and conservative think tanks. And infer ideology based on how similar an outlet's citation patterns are to legislators'. But such measures are sensitive to the dates, and the specific think tanks included (Gasper 2011).

Others use information in editorials. They classify newspapers based on which party's candidate and which Supreme Court decision they endorse (Ho, Quinn et al. 2008; Puglisi and Snyder 2015). But editorials are a small section of the newspapers. And editorial content may not resemble the overall tone of the newspaper. For example, WSJ's editorial section is well to the right of its news section (Barberá and Sood 2014).

Yet others exploit the variation in coverage of facts that reflect badly on a party (Larcinese, Puglisi and Snyder 2011). If a media organization is more likely to cover scandals of a particular party, it is likely biased against that party. But this kind of inference runs the risk of maligning all coverage with biases observed in small portions of the coverage (Budak, Goel and Rao 2016).

Lastly, we can use speeches made on the Congressional floor to estimate the relationship between words and ideology. And then use the model to infer ideology based on words used in the news (Gentzkow and Shapiro 2010; Sood and Guess 2015). There are at least three problems with such measures. First, the words used by members of Congress aren't the same as used by the news. So the model may not be valid when predicting ideology based on words used in the news. Second, performance within training data is poor. For instance, the r^2 of some of the published models are no greater than .7 (see, for e.g., Barberá and Sood 2014; Gentzkow and Shapiro 2010). And within party r^2 of models is generally only half as large (Barberá and Sood 2014). Third, the method makes strong judgments about the overlap in time between congressional speech and news. This, in turn, relies on assumptions about overlap between congressional and the news media agenda.

Lastly, some people provide explicit directions to casual workers on micro-task markets like Amazon's Mechanical Turk (Garz et al. 2017; Budak, Goel and Rao 2016; Baum and

Groeling 2008). The quality of coding depends on incentives and clarity of directions. The intercoder reliability in Budak, Goel and Rao (2016); Baum and Groeling (2008), for instance, is 81%. This suggests a fairly healthy dose of measurement error, even if one were to assume zero bias. In all, measurement of ideology remains a challenge.

Measurement of Consequences

If our aim is to find a behavioral analog to preferences, ‘selectivity’ seems reasonable. To estimate consequences of media diets, we need to know how many calories came from which item. For instance, if we want to explain political knowledge, the pertinent variable is the total amount of correct political information consumed. If it is misinformation, the variable is total incorrect political information. For explaining knowledge about a topic, we need to tally all correct information about that topic consumed by a person. And if the dependent variable is partisan affect, we may want two separate totals. We may want total congenial information consumed. And given exposure to uncongenial information likely also polarizes, we may also want to tally consumption of uncongenial information.

Studies estimating the causal impact of interest in consuming political information are hard to come by. One reason for the shortfall is political interest is largely stable (Prior 2010). Even so, interest in political campaigns does increase over the course of the campaign (NAES File). And exogenous factors like the closeness of the race likely matter too (Gerber and Green 2000).

But most work takes preferences as fixed and estimates the impact of changes in the information environment. Snyder Jr and Strömberg (2008), for instance, identify the impact of local newspapers on knowledge about local representatives by using discontinuities in newspaper distribution areas. Others exploit laws affecting the price of broadband (Lelkes, Sood and Iyengar 2015). Yet others exploit exogenous changes in menus due to entry and exit of channels (Martin and Yurukoglu 2014). And yet others exploit cross-national variation in in-

stitutions (Lelkes 2016; Curran et al. 2009). Curran et al. (2009) assume that questions about international politics are equally hard for everyone. And infer that Americans know less than Danes and the Fins. But attributing political ignorance to media “systems” is hard, near impossible. The source of variation could be institutions, “culture”, education systems, or weather. Deducting knowledge of civics may help rule out some of the explanations.

But more problematically, all these research designs give us a brief and shallow glimpse into consequences of imbalances in media exposure. For instance, we never get to observe, not with any great precision, how much do people already consume. This means that we must elide over important variation.

Another important limitation is that with instrumental variable models, we only get ‘Local Average Treatment Effect’ (LATE) estimates. For instance, in Lelkes, Sood and Iyengar (2015), if the assumptions hold, the effect is driven by people who were encouraged to adopt broadband. The adopters, however, may not be of sufficient theoretical or political interest. And changes among adopters do not tell us what would happen if the entire population were to get broadband. If the treatment effects vary by kinds of people, as they surely do, the population estimates can be far from LATE. Garz (2017) offers another example. It estimates effect of retirement on knowledge. But extra leisure time at 65 is very likely to be spent differently than when someone is 21.

The impacts of changes in the information environment over time on consumption have been harder to pin down. One reason for that is identifying changes in consumption in the presence of substitutes and variation in quality is hard. So researchers have generally sidestepped it.

Since the menus from which people choose in the real world cannot be manipulated easily, lab studies have taken the stage. Manipulations in laboratory or online survey experiments, however, lack ecological validity. The menu of choices in most survey experiments is exceedingly small. And the relative change in the menu as a result of changing even one item,

large. These large changes don't translate well to the real world, where the menu is vast, if not limitless, though people limit most of their consumption to a small set of sources, and many implicitly whittle down choices to a manageable few.

A more significant problem with lab studies is the identification of media effects using forced exposure. Studies providing limited choice suggest that forced exposure exaggerates the impact of exposure to partisan news (Arceneaux, Johnson and Murphy 2012; Arceneaux and Johnson 2013). Those disinterested in news are unlikely to see it, while those who prefer news tend to hold strong attitudes, and are thus immune to its effects.

A yet more important limitation of media effects research is that the treatment effects are often not calibrated. If we ever want to generalize the findings, we need to know the exact dose delivered. There are a few exceptions. Curran et al., for instance, measure informativeness. But, the treatments are notoriously multidimensional. And we still struggle to map what aspect of the treatment, in what quantity, explains the changes we see. ,

Discussion

The media have been revolutionized—repeatedly. In the 1970s, most of us wouldn't have forecast the success of cable, and the overwhelming number of options available on it. In the 1980s, most wouldn't have forecast the current version of the Internet. And in the early 2000s, most of us wouldn't have forecast the current smartphones, and their success. Or the social media and its success. And again only a few of us would have forecast that miniaturization of cameras and social media would come together to highlight police malpractice. For even if we boast of understanding the larger capitalist institution behind the production of technology and content, the future constantly surprises us. For not only do we not understand how technology is created and appropriated, we do not understand *us* very well.

As social scientists, our work is cut out. And while the tools at our disposal have never

been better—access to massive amounts of data, super fast computers along with incredible software that allows us to not only calculate complex things but also allows us to fail quickly, the challenge has also never been greater—the number of devices per person on which people consume information is expected to exceed 6 in the next 5 years.

Add to this, the pessimism that many media scholars feel. The move away from media carried over public waves means that today the American government has few levers with which to influence what media outlets can carry. And in the absence of expected policy consequences, some media scholars wonder about the purpose of research. But the relationship between research and policy has always been regrettably weak, even in areas where the government enjoys great policy-making capacity. Instead, research often finds its ways via products. For instance, measures of media ideology can be used to build media consumption tools that alert people about the potential bias.

In all, social scientists have taken some decisive steps in quantifying the extent to which people consume various media, but a variety of issues remain. In this chapter, we have tried to clarify what issues impinge upon our inferences to what extent. We also tried to make a case that some more original thinking is needed in how we conceptualize the independent variable and the potential dependent variables. In particular, we think that net imbalance rather than ‘selectivity’ is a more apt variable for many of the decisions. More generally, we hope that the discussion in the chapter provokes greater debate on measurement, and provides greater clarity about the trade-offs.

References

- Akerlof, George A and Robert J Shiller. 2015. *Phishing for phools: The economics of manipulation and deception*. Princeton University Press.
- Arango-Kure, Maria, Marcel Garz and Armin Rott. 2014. “Bad news sells: The demand for news magazines and the tone of their covers.” *Journal of Media Economics* 27(4):199–214.
- Arceneaux, Kevin and Martin Johnson. 2013. *Changing Minds Or Changing Channels?: Partisan News in an Age of Choice*. University of Chicago Press.
- Arceneaux, Kevin, Martin Johnson and Chad Murphy. 2012. “Polarized political communication, oppositional media hostility, and selective exposure.” *The Journal of Politics* 74(01):174–186.
- Bakshy, Eytan, Solomon Messing and Lada Adamic. 2015. “Exposure to ideologically diverse news and opinion on Facebook.” *Science* p. aaa1160.
- Barberá, Pablo and Gaurav Sood. 2014. “Follow Your Ideology: A Measure of Ideological Location of Media Sources.”
- Bartels, Larry M. 2005. “Homer gets a tax cut: Inequality and public policy in the American mind.” *Perspectives on Politics* 3(1):15–31.
- Baum, Matthew A and Tim Groeling. 2008. “New media and the polarization of American political discourse.” *Political Communication* 25(4):345–365.
- Bawn, Kathleen, Martin Cohen, David Karol, Seth Masket, Hans Noel and John Zaller. 2012. “A theory of political parties: Groups, policy demands and nominations in American politics.” *Perspectives on Politics* 10(3):571–597.

- Bennett, W Lance and Shanto Iyengar. 2008. "A new era of minimal effects? The changing foundations of political communication." *Journal of Communication* 58(4):707–731.
- Budak, Ceren, Sharad Goel and Justin M Rao. 2016. "Fair and balanced? quantifying media bias through crowdsourced content analysis." *Public Opinion Quarterly* 80(S1):250–271.
- Chan, J. and W. Suen. 2008. "A spatial theory of news consumption and electoral competition." *Review of Economic Studies* 75(3):699–728.
- Curran, James, Shanto Iyengar, Anker Brink Lund and Inka Salovaara-Moring. 2009. "Media System, Public Knowledge and Democracy A Comparative Study." *European Journal of Communication* 24(1):5–26.
- DellaVigna, Stefano and Ethan Kaplan. 2007. "The Fox News Effect: Media Bias and Voting." *The Quarterly Journal of Economics* 122(3):1187–1234.
- DellaVigna, Stefano, John A List and Ulrike Malmendier. 2012. "Testing for altruism and social pressure in charitable giving." *The quarterly journal of economics* 127(1):1–56.
- Dilliplane, Susanna. 2011. "All the news you want to hear: The impact of partisan news exposure on political participation." *Public Opinion Quarterly* 75(2):287–316.
- Dilliplane, Susanna, Seth K Goldman and Diana C Mutz. 2013. "Televised exposure to politics: New measures for a fragmented media environment." *American Journal of Political Science* 57(1):236–248.
- Druckman, James N, Matthew S Levendusky and Audrey McLain. 2015. "No Need to Watch: How the Effects of Partisan Media Can Spread via Interpersonal Discussions." *American Journal of Political Science* .
- Festinger, Leon. 1957. *A theory of cognitive dissonance*. Stanford University Press.

- Flaxman, Seth, Sharad Goel and Justin M Rao. 2016. "Filter bubbles, echo chambers, and online news consumption." *Public Opinion Quarterly* 80(S1):298–320.
- Garrett, R. Kelly. 2009. "Politically motivated reinforcement seeking: Reframing the selective exposure debate." *Journal of Communication* 59(4):676–699.
- Garz, Marcel. 2017. "Retirement, consumption of political information, and political knowledge." *European Journal of Political Economy* .
- Garz, Marcel, Gaurav Sood, Daniel F Stone and Justin Wallace. 2017. "What Drives Demand for Media Slant?"
- Gaspar, John T. 2011. "Research Note Shifting Ideologies? Re-examining Media Bias." *Quarterly Journal of Political Science* 6:85–102.
- Genova, Bistravapka KL and Bradley S Greenberg. 1979. "Interests in news and the knowledge gap." *Public Opinion Quarterly* 43(1):79–91.
- Gentzkow, Matthew. 2006. "Television and voter turnout." *The Quarterly Journal of Economics* pp. 931–972.
- Gentzkow, Matthew and Jesse M Shapiro. 2006. "Media bias and reputation." *Journal of Political Economy* 114(2):280–316.
- Gentzkow, Matthew and Jesse M Shapiro. 2010. "What drives media slant? Evidence from US daily newspapers." *Econometrica* 78(1):35–71.
- Gentzkow, Matthew and Jesse M. Shapiro. 2011. "Ideological Segregation Online and Offline." *The Quarterly Journal of Economics* 126(4):1799–1839.
- Gerber, Alan S and Donald P Green. 2000. "The effects of canvassing, telephone calls, and direct mail on voter turnout: A field experiment." *American political science review* 94(3):653–663.

- Goldman, Seth K and Diana C Mutz. 2011. "The friendly media phenomenon: A cross-national analysis of cross-cutting exposure." *Political Communication* 28(1):42–66.
- Graber, Doris. 2004. "Mediated politics and citizenship in the twenty-first century." *Annu. Rev. Psychol.* 55:545–571.
- Groseclose, Tim and Jeffrey Milyo. 2005. "A measure of media bias." *The Quarterly Journal of Economics* 120(4):1191–1237.
- Guess, AM. 2016. "Media choice and moderation: Evidence from online tracking data." *Unpublished manuscript, New York University* .
- Hindman, Matthew. 2008. *The myth of digital democracy*. Princeton University Press.
- Ho, Daniel E, Kevin M Quinn et al. 2008. "Measuring explicit political positions of media." *Quarterly Journal of Political Science* 3(4):353–377.
- Iyengar, Shanto and Kyu S. Hahn. 2008. "Red media, blue media: Evidence of ideological selectivity in media use." *Journal of Communication* 59(1):19–39.
- Iyengar, Shanto and Kyu S Hahn. 2009. "Red media, blue media: Evidence of ideological selectivity in media use." *Journal of Communication* 59(1):19–39.
- Iyengar, Shanto, Kyu S. Hahn, Jon A. Krosnick and John Walker. 2008. "Selective Exposure to Campaign Communication: The Role of Anticipated Agreement and Issue Public Membership." *The Journal of Politics* 70(01):186–200.
- Johnson, Valen E and James H Albert. 2006. *Ordinal data modeling*. Springer Science & Business Media.
- Katz, Elihu. 1957. "The two-step flow of communication: An up-to-date report on an hypothesis." *Public opinion quarterly* 21(1):61–78.

- Keeter, Scott and Harry Wilson. 1986. "Natural treatment and control settings for research on the effects of television." *Communication Research* .
- Knobloch-Westerwick, Silvia and Jingbo Meng. 2009. "Looking the other way selective exposure to attitude-consistent and counterattitudinal political information." *Communication Research* 36(3):426–448.
- Knobloch-Westerwick, Silvia and Jingbo Meng. 2011. "Reinforcement of the political self through selective exposure to political messages." *Journal of Communication* 61(2):349–368.
- Kohut, Andrew, Richard Morin and Scott Keeter. 2007. "What Americans Know: 1989-2007, Public Knowledge of Current Affairs Little Changed by News and Information Revolutions." *PEW Research Center*, April 15.
- Krugman, Herbert E and Eugene L Hartley. 1970. "Passive learning from television." *Public Opinion Quarterly* 34(2):184–190.
- Larcinese, Valentino, Riccardo Puglisi and James M Snyder. 2011. "Partisan bias in economic news: Evidence on the agenda-setting behavior of US newspapers." *Journal of Public Economics* 95(9):1178–1189.
- Leighley, Jan E. 2003. *Mass media and politics: A social science perspective*. Houghton Mifflin College Division.
- Lelkes, Yphtach. 2016. "Winners, losers, and the press: The relationship between political parallelism and the legitimacy gap." *Political Communication* 33(4):523–543.
URL: <http://dx.doi.org/10.1080/10584609.2015.1117031>
- Lelkes, Yphtach. 2017. "Mind the time : The implications of high-speed internet for political behavior." *Working Paper* .

- Lelkes, Yphtach, Gaurav Sood and Shanto Iyengar. 2015. "The Hostile Audience: The Effect of Access to Broadband Internet on Partisan Affect." *American Journal of Political Science* .
- Levendusky, Matthew S. 2013. "Partisan Media Exposure and Attitudes Toward the Opposition." *Political Communication* 30(4):565–581.
- Lord, Charles G, Lee Ross and Mark R Lepper. 1979. "Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence." *Journal of Personality and Social Psychology* 37(11):2098–2109.
- Luskin, Robert C. 1990. "Explaining political sophistication." *Political Behavior* 12(4):331–361.
- Luskin, Robert C, James S Fishkin and Roger Jowell. 2002. "Considered opinions: Deliberative polling in Britain." *British Journal of Political Science* 32(3):455–487.
- Martin, Gregory J and Ali Yurukoglu. 2014. "Bias in Cable News: Real Effects and Polarization."
- Melican, Debra Burns and Travis L Dixon. 2008. "News on the net credibility, selective exposure, and racial prejudice." *Communication Research* 35(2):151–168.
- Metzger, Miriam J, Ethan H Hartsell and Andrew J Flanagin. 2015. "Cognitive Dissonance or Credibility?: A Comparison of Two Theoretical Explanations for Selective Exposure to Partisan News ." *Communication Research* .
URL: <http://crx.sagepub.com/content/early/2015/11/24/0093650215613136.abstract>
- Morris, Jonathan S. 2007. "Slanted Objectivity? Perceived Media Bias, Cable News Exposure, and Political Attitudes*." *Social Science Quarterly* 88(3):707–728.
- Prior, Markus. 2003. "Any good news in soft news? The impact of soft news preference on political knowledge." *Political Communication* 20(2):149–171.

- Prior, Markus. 2005. "News vs. entertainment: How increasing media choice widens gaps in political knowledge and turnout." *American Journal of Political Science* 49(3):577–592.
- Prior, Markus. 2007. *Post-broadcast democracy: How media choice increases inequality in political involvement and polarizes elections*. Cambridge University Press.
- Prior, Markus. 2009. "The immensely inflated news audience: Assessing bias in self-reported news exposure." *Public Opinion Quarterly* 73(1):130–143.
- Prior, Markus. 2010. "Youâ€™ve either got it or you donâ€™t? The stability of political interest over the life cycle." *The Journal of Politics* 72(03):747–766.
- Prior, Markus. 2013. "Media and political polarization." *Annual Review of Political Science* 16:101–127.
- Puglisi, Riccardo and James M Snyder. 2015. "The balanced US press." *Journal of the European Economic Association* 13(2):240–264.
- Robinson, Michael J. 1976. "Public Affairs Television and the Growth of Political Malaise: The Case of â€”The Selling of the Pentagonâ€”." *American political science review* 70(02):409–432.
- Rosenthal, Robert and Ralph L Rosnow. 2009. *Artifacts in Behavioral Research: Robert Rosenthal and Ralph L. Rosnow’s Classic Books*. Oxford University Press.
- Samejima, Fumiko. 1969. "Estimation of latent ability using a response pattern of graded scores." *Psychometrika monograph supplement* .
- Schudson, Michael. 1998. *The good citizen: A history of American civic life*. Free Press.
- Sears, David O and Jonathan L Freedman. 1963. Commitment, information utility, and selective exposure. Technical report CALIFORNIA UNIV LOS ANGELES.

- Sears, David O and Jonathan L Freedman. 1967. "Selective exposure to information: A critical review." *Public Opinion Quarterly* 31(2):194–213.
- Shanahan, James and Michael Morgan. 1999. *Television and its viewers: Cultivation theory and research*. Cambridge University Press.
- Shineman, Victoria. 2012. Incentivizing Participation Increases Political Information: Evidence from a Randomized Field Experiment. In *APSA 2012 Annual Meeting Paper*.
- Snyder Jr, James M and David Strömberg. 2008. Press coverage and political accountability. Technical report National Bureau of Economic Research.
- Sood, Gaurav. 2014. How Can You Think That?: Deliberation and the Learning of Opposing Arguments. In *Annual Meeting of the American Political Science Association*.
- Sood, Gaurav and Andy Guess. 2015. Measures of Ideology: Agendas, and Positions on Agendas. In *New Directions in Text as Data Analysis*.
- Sood, Gaurav and Daniel Trielli. 2016. The Face of Crime in Prime Time: Evidence from Law and Order. Technical report.
- Sood, Gaurav and Shanto Iyengar. 2013. All in the Eye of the Beholder: Partisan Affect and Ideological Accountability. In *Annual Meeting of the American Political Science Association*.
- Sørensen, Rune J. 2016. "The Impact of State Television on Voter Turnout." *British Journal of Political Science* (2006):1–22.
URL: http://www.journals.cambridge.org/abstract_S000712341600048X
- Stroud, Natalie Jomini. 2010. "Polarization and partisan selective exposure." *Journal of Communication* 60(3):556–576.

- Trussler, Marc and Stuart Soroka. 2014. "Consumer demand for cynical and negative news frames." *The International Journal of Press/Politics* 19(3):360–379.
- Tversky, Amos and Daniel Kahneman. 1971. "Belief in the law of small numbers." *Psychological bulletin* 76(2):105.
- Vallone, Robert P, Lee Ross and Mark R Lepper. 1985. "The hostile media phenomenon: biased perception and perceptions of media bias in coverage of the Beirut massacre." *Journal of personality and social psychology* 49(3):577.
- Van Kempen, Hetty. 2007. "Media-party parallelism and its effects: A cross-national comparative study." *Political Communication* 24(3):303–320.
- Zaller, John. 1992. *The nature and origins of mass opinion*. Cambridge university press.
- Zaller, John. 2003. "A new standard of news quality: Burglar alarms for the monitorial citizen." *Political Communication* 20(2):109–130.
- Zukin, Cliff and Robin Snyder. 1984. "Passive learning: When the media environment is the message." *Public Opinion Quarterly* 48(3):629–638.

Appendix A American National Election Studies

The data are from the American National Election Studies Cumulative Data File.

Appendix A.1 Political Knowledge Over the Lifetime

The data are from the American National Election Studies. And each line represents data from a different year. Political knowledge was assessed by the interviewer.

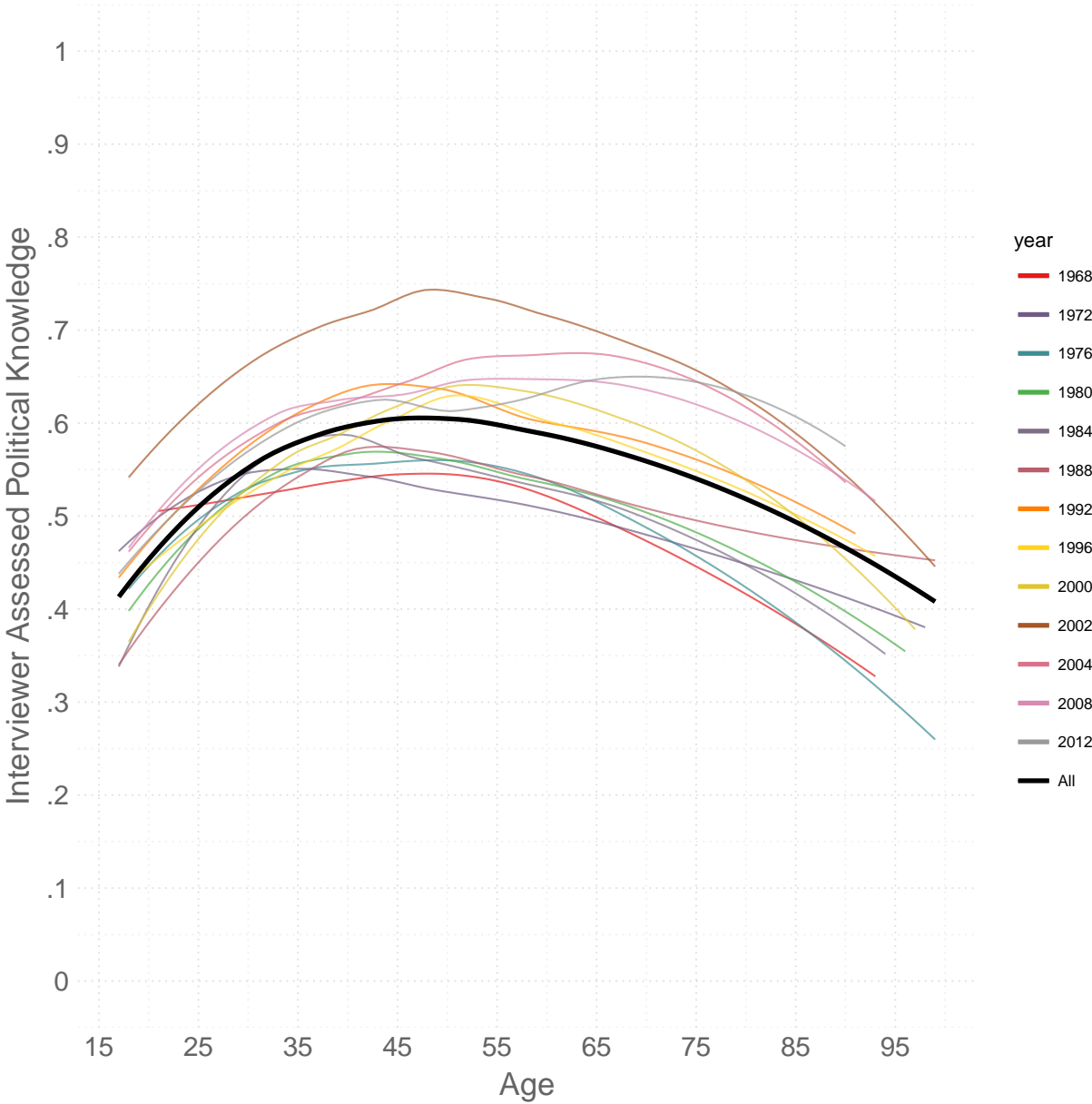
Measures

Political interest (VCF0313) typically ran as follows: Some people seem to follow what's going on in government and public affairs most of the time, whether there's an election going on or not. Others aren't that interested. Would you say you follow what's going on in government and public affairs most of the time, some of the time, only now and then, or hardly at all?"

Education (VCF0140a) measures education in 7 categories: 8 grades or less ('grade school'), 9-12 grades ('high school'), no diploma/equivalency, 12 grades, diploma or equivalency, 12 grades, diploma or equivalency plus non-academic training, Some college, no degree; junior/community college level degree (AA degree), BA level degrees, Advanced degrees incl.

LLB

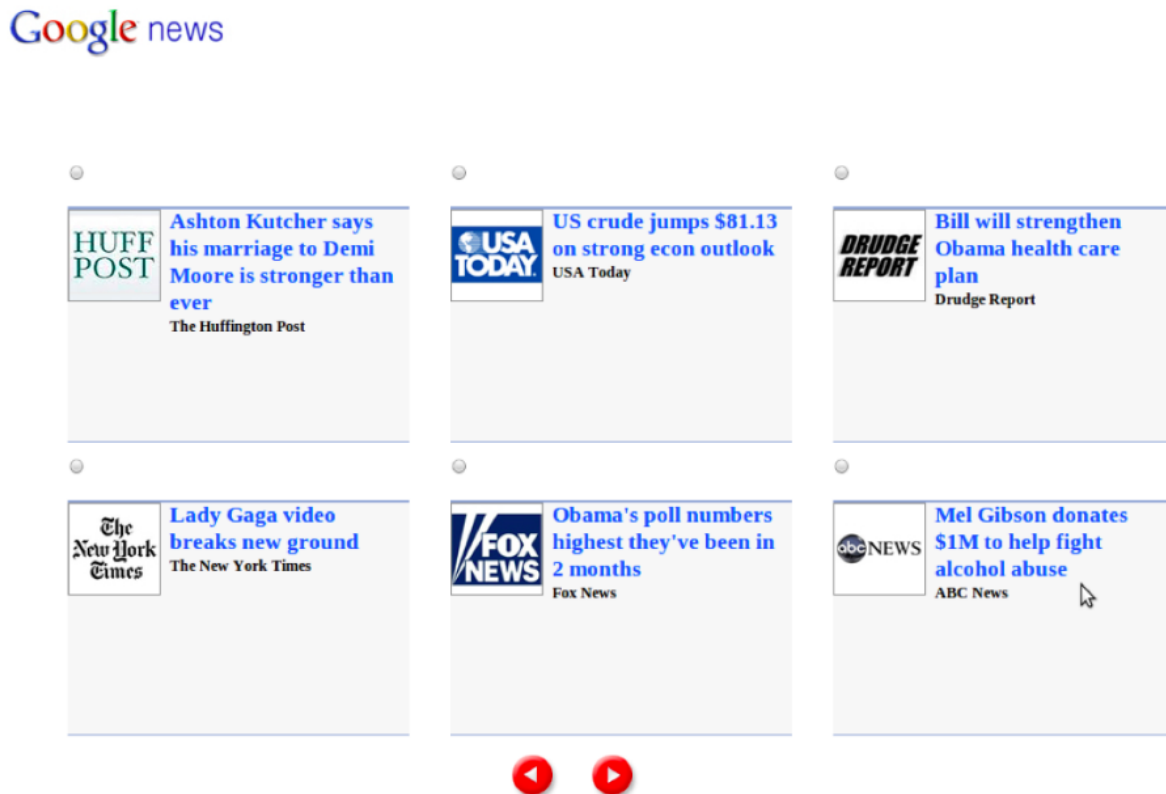
Figure A1: Political Knowledge Over the Lifetime



Appendix B YouGov Study

Appendix B.1 Measurement of Preference for Hard News and News from Congenial Sources

Figure B1: News Organization Trial

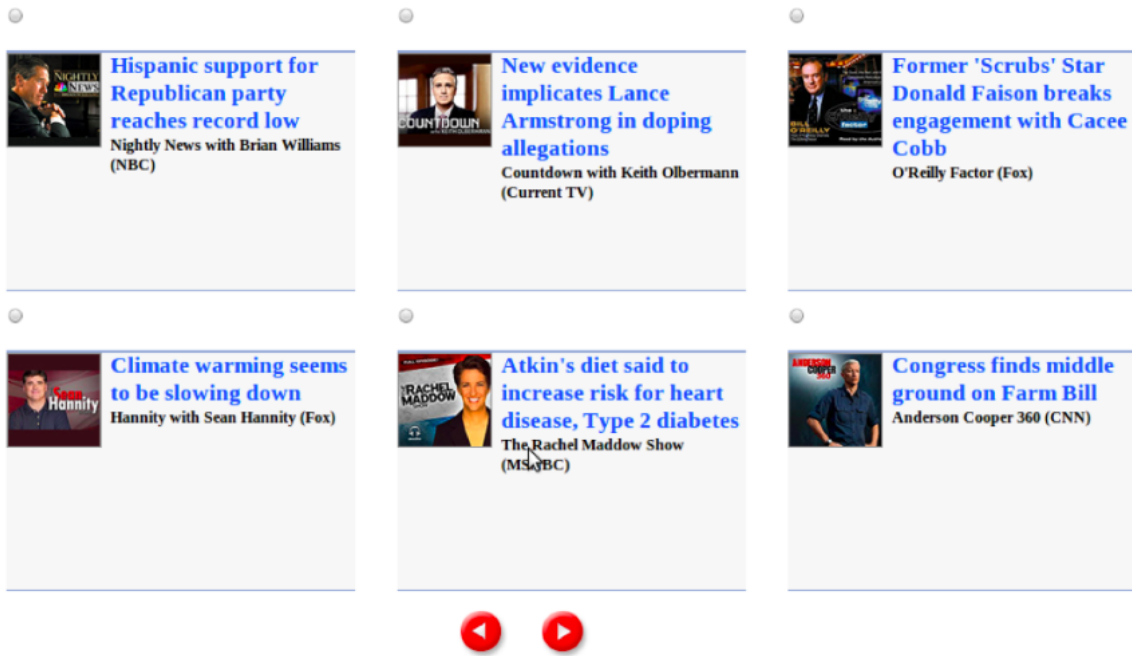


Following the four news selection screens, we asked respondents to select a short video sound bite from major political figures. We had included the video selection task to assess the overlap in selectivity between news sources and campaign advertising, and to investigate whether people who preferred to get their news from partisan sources also preferred to listen to their party's candidates engage in attack-oriented rather than promotional rhetoric.

There were two video selection screens. Each screen featured one prominent Democrat and one prominent Republican making either a positive or negative appeal. For example,

Figure B2: News Personality Trial

Google news



respondents could select between two Romney clips, one titled “I’ll get America working again” and another titled “We have zero faith in our President.” While the first screen presented clips from Obama and Romney, the second screen included clips from Nancy Pelosi and Michele Bachmann. The order of videos within each of the two screens was randomized.

Preference for Soft News. If a respondent selected a soft news story on a trial, we coded it as 1. We coded hard news selections as 0. We next fitted the choices as a function of an underlying latent trait (preference for soft news) using a two-parameter item response model (Lord, Ross and Lepper 1979). Letting i index items (trials), j index respondents, the model can be written as follows:

$$P(y_{ij} = 1|\eta_j) = \frac{\exp(\beta_i + \lambda_i\eta_j)}{1 + \exp(\beta_i + \lambda_i\eta_j)} \quad (1)$$

In Equation 1, y_{ij} is respondent j ’s response to item i , η_j is the unobserved preference for soft news for respondent j , λ_i is the discrimination parameter telling us how much responses to item (trial) i distinguish between respondents who prefer soft news more from those to who prefer it less, and $\frac{-\beta_i}{\lambda_i}$ is the “difficulty” parameter. We recover the ordinal arrangement of respondents on the latent dimension. We rescale the scores to lie between 0 (weakest preference for soft news) to 1 (strongest preference for soft news).

Preference for Partisan Congenial Sources. To estimate preference for news from partisan congenial sources, we categorized the items into three ordinal categories—left-leaning, right-leaning and non-partisan. We then fit the participants’ choices to a graded response model (Samejima 1969; Johnson and Albert 2006), which can be seen as an extension of the two-parameter logistic item response model that we described above.

We rescaled the estimate of preference for partisan congenial sources to lie between 0 (strongest preference for left-leaning sources in our data) and 1 (indicating the strongest preference for right-leaning sources in our data).

Scandal Knowledge: To assess knowledge of the IRS scandal, we asked the respondents to identify which groups received extra scrutiny by the IRS (right-wing groups), the name of the IRS commissioner that resigned (Steve Miller), and the location of the IRS office at the center of the scandal (Cincinnati). We added the number of correct answers and rescaled the sum to lie between 0 and 1.

Feelings Toward Parties: In the survey, we asked respondents about their feelings towards the parties.

Appendix B.2 Correlation Between Preference for Information from Congenial Sources and Partisan Affect and Scandal Knowledge

People with the strongest preference for partisan media are roughly 7% more polarized than those who prefer moderate sources or indiscriminate (see left panel of Figure B3).

We also expect exposure to partisan congenial media to skew what people know. Partisan media cover congenial information more often than uncongenial information (Larcinese, Puglisi and Snyder 2011). And such biases can lead to partisan gaps in knowledge and misinformation. Such biases may be responsible for partisans believing that the leaders of their own party are less extreme than they are (Sood and Iyengar 2013).

To shed light on the issue, we collected data to assess the extent of selective learning. Roughly one year after the initial YouGov study, we surveyed 1,000 respondents in a second wave. Between the first and second wave of our study, the Obama administration was engulfed in the IRS scandal. Different media outlets . Searching archive.org's TV News database for 'IRS' AND 'Tea Party' from January until July 2013 reveals 806 mentions on Fox News, 464 mentions on MSNBC, and 266 mentions on CNN. Given this asymmetry, we expected an asymmetry in knowledge of the IRS scandal between those that preferred partisan media to non-partisan media. And since, right-wing partisan media likely devoted more attention to the IRS scandal

than did the left-wing media, we expected Republicans with strong preferences for partisan sources to be more informed about the scandal than Democrats with similar preferences. The expectation was born out (right panel of Figure B3). Those with the strongest preferences for right-wing media knew roughly 10 percent more about the IRS scandal than those with the strongest preferences for left-wing media and those with ecumenical preferences.

Figure B3: The Relationship between Preference for News from Congenial Sources and Affective Polarization & Scandal Knowledge

